Adaptive Trend Analysis - A Simple Solution to Data Variability

Larry A. Toms
Larry A. Toms Technical Services
5018 High Pointe Drive
Pensacola, FL 32505-1825

1998

Abstract: *This paper discusses variability problems with oil test data. Simple adaptive trend algorithms can be used to reduce the effect of data variability (from sampling, testing & related maintenance procedures) on trend analysis. The method described reduces the uncertainty in sample interpretation results. Examples are used to emphasize key points.*

Oil Data Variability: Making a reliable recommendation with specific maintenance instructions from routine sample data is a necessary requirement for a condition monitoring program. However, reliable interpretation is often difficult because of a normally high variability in sample data. Trend calculations are easily compromised by maintenance, sampling and testing procedures. The following are examples of the more common problems with associated general purpose solutions, which can dramatically improve data trending reliability.

Effect of High and Low Readers: The nominal test readings from two machines of the same type, operating in the same environment are usually different. Sample readings usually range from a nominal low value to a nominal high value requiring different baselines for trend line analysis for the individual machines. Consider the wear trends shown in Figure 1, the readings of the first four samples (Equipment A and Equipment B) show the typical nominal range. When using simple condemning alarms, the same increase in the wear rate would be interpreted differently for Equipment A then for Equipment B.
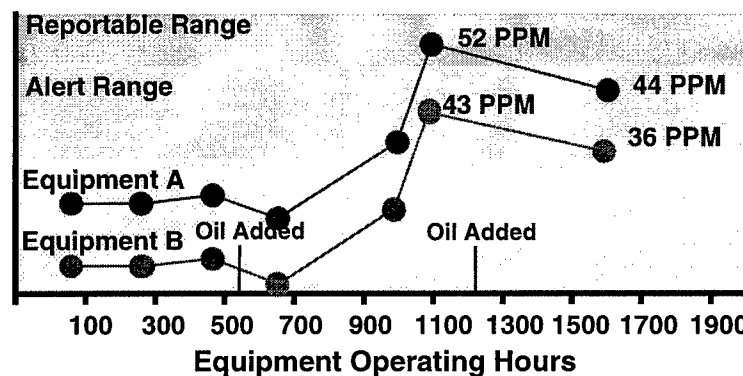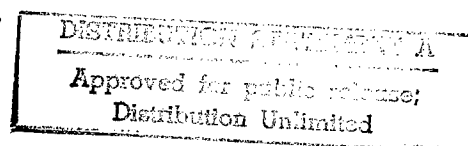

Figure 1: Trends of Similar Equipment in Same Service

Few oil analysts would misinterpret the data shown in Figure 1, but in the real world, single machine data is not comparatively displayed with idealized data or data from other machines from the same equipment family. The dependence on simple limits often results in misinterpretation.

Effect of Data Measurement Variability: Consider the spectrometric data presented in Figure 2. This table shows statistics calculated from six successive weeks of sample data (~4500 samples) generated by a rotary disk emission (RDE) spectrometer from a fleet of 1500 diesel engines; followed by an additional six successive weeks of sample data (~4500 samples), from the same fleet, generated by an inductively coupled plasma (ICP) spectrometer.

RDE Spectrometer	Pb	Cu	Sn	Fe	Cr	Al	B	Na	Si
Average:	16.2	20.1	2.0	32.0	8.5	4.3	5.2	27.2	8.9
Standard Deviation:	5.5	4.3	1.7	5.7	4.0	1.8	5.4	19.8	4.1
ICP Spectrometer									
Average:	5.0	8.0	1.1	15.9	1.5	2.4	2.4	17.8	6.5
Standard Deviation:	2.8	2.4	0.9	4.2	1.5	0.5	2.6	16.3	2.9

Figure 2: Differences between Tests on the Same Equipment Fleet (~4500 Samples)

Note the very large differences in readings between the RDE and ICP instruments! Different oil laboratories may use different instruments, and larger labs may have several different instrument models. Do you know which instrument type your samples were analyzed on? Or, if samples are always analyzed on the same instrument? Samples analyzed on different instruments often give different results, even when the instrument is of the same type, but of different manufacturer—this is particularly true if the instruments are not enrolled in a correlation program. Nor is the problem limited to spectroscopy, oil physical property measurements can vary significantly from instrument to instrument and operator to operator using the same instrument! In addition, many of the common oil property tests are operator sensitive and different technicians often generate different results.

Effect of Maintenance and Sampling Practice: Consider the trend plot shown in Figure 3. The last two samples (**A1 & A2**) were taken at intervals longer than the specified sample interval. When this happens, the data lost often masks the effect of maintenance activities and reduces potential warning time available to plan a maintenance response in the event the last sample is abnormal. The second last sample (**A1**) does not indicate the high rate of change (trend slope) that would have been indicated if sample **C** was taken at 1350 hours. Nor does it indicate the actual effect of the oil addition (1400 hours) that would have been indicated by the missed sample (**D**) at 1500 hours. The long interval for the last sample (**A2**) further reduced the effective rate of change indication and although it results in an alarm, significant warning time was lost. The missed sample **E** (1650 hours) would have restored some indication that a high trend was present, however the combination of all missed samples (**C, D & E**) in conjunction with the oil addition at 1400 hours effectively eliminated usable trend history until the last sample (**A2**).
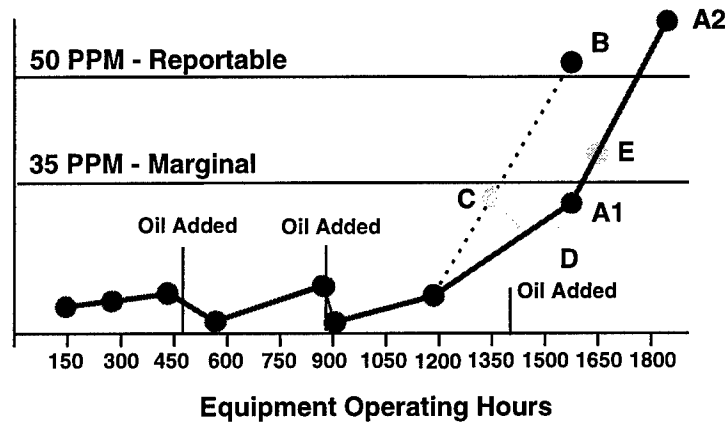
Figure 3: Effect of Long Sample Intervals

Note, had the samples at 1350 & 1500 hours (**C & B**) been taken on time and the oil not added at 1400 hours, the warning trend would have occurred over 300 hours earlier.

Oil Data Interpretation Solutions: The most important solution for data variability due to sampling, testing or maintenance procedures is to provide *management oversight to ensure all procedures are carried out properly*. While different corporate departments and/or commercial contractors make the task of quality assurance difficult to manage; accepted certification standards, such as ISO 9002, can be applied to obtain effective results.

The second most important factor in oil data variability is the proper scheduling of sampling and oil related equipment maintenance. For example, samples taken after an oil change provide a good indication that the oil was changed but no indication of oil or machine condition—that data was disposed of with the oil. Oil, filter and component changes should be scheduled after sampling to preserve condition history. The exception to this rule occurs when the oil change and sample interval are equal. The data from a sample taken just prior to oil drain will always show the oil at the end of its usable life—a fact already apparent by scheduled oil change—unless the oil is being changed prematurely. When the optimum sample interval is effectively equal to the oil change interval, the actual sample interval should be set to get at least two samples per oil charge. One at about 30% to 50% of the oil drain interval and the other just prior to drain. Data from these two sample points and new oil samples will significantly improve trend analysis results.

After sampling, testing and maintenance procedures are improved to the extent possible, mathematical data interpretation procedures may be used to reduce the effect of any remaining data variability. However, some scheduling conflicts will continue to occur and data trending and interpretation procedures must be adaptive to overcome these interferences. An adaptive trend system requires some standards to permit the system to distinguish between normal and abnormal procedures. Normalization can be established

by assigning a standard sample interval for each machine type from which the trends will be calculated. The most appropriate trend equation may be selected in accordance with oil change and sample interval data to determine the most reliable trend data.

Rise over run from previous sample: For routine samples that are taken between one-half and one and one-half (0.5 to 1.5) the standard sample interval—providing the oil has not been changed since the last sample. Figure 4 shows a *rise-over-run* equation which will calculate a reliable trend accounting for the actual equipment usage.

$$Trend = \left(\frac{CS - LS}{AI}\right) * SI$$

CS: current sample data,
LS: last sample data,
AI: actual time-on-oil since the last sample, and;
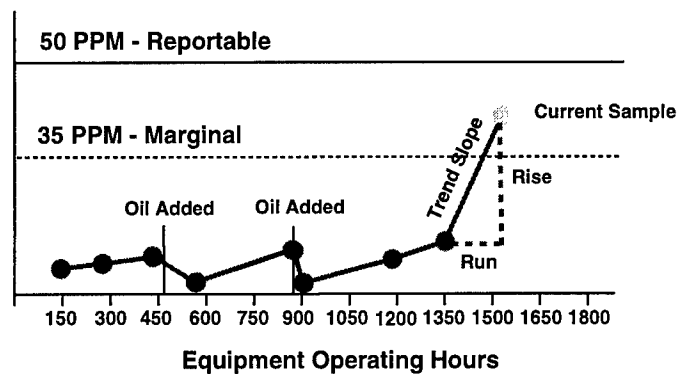**SI:** established interval for equipment type.

Figure 4: Rise over Run Calculation

Rise over run from a statistically derived historical point: If the sample *interval is smaller than 0.5 times the standard interval* and there are sufficient samples, use a linear regression to predict the value of a sample one standard interval prior the current sample. Likewise, if the sample *interval is greater than 1.5 times the standard interval* and there are sufficient samples, use a linear regression to predict the value of a sample one standard interval prior the current sample. Trend data can then be calculated by a rise-over-run equation using the predicted sample and current sample data.

$$Trend = \left(\frac{CS - PS}{SI}\right) * SI$$

CS: current sample data,
PS: predicted last sample data, and;
SI: established interval for equipment type.

While more sophisticated nonlinear or curve fitting techniques can be used, the author has found this method to be satisfactory for automated oil analysis trending. In Figure 5, note the difference in trend slope for the predicted last sample versus the trend slope from the actual last sample. Long sample intervals tend to reduce trend values and alarm indications.
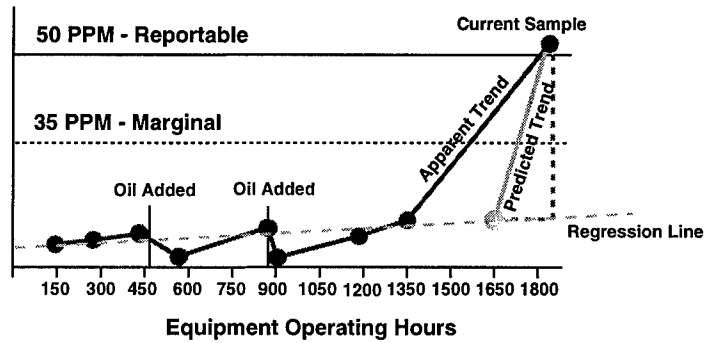
Figure 5: Calculating Trends from a Predicted Last Sample

Short sample intervals, such as occurs when check samples are taken (Figure 6), can effectively destroy the actual trend because of the short interval involved. Again, a regression line can be used to predict a reasonable last sample value, dramatically improving the rise over run trend calculation.
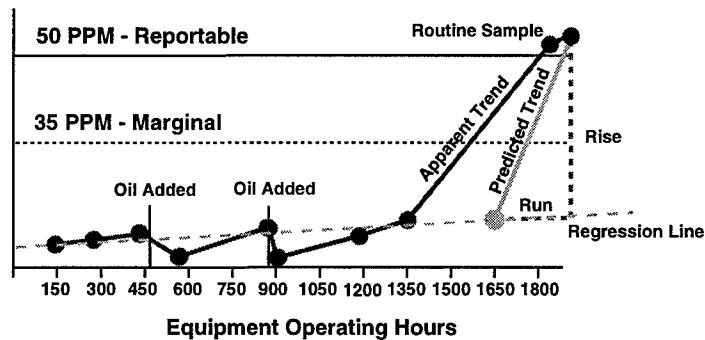

Figure 6: Calculating Trends for a Check Sample

Rise over run from nominal value or a nominal interval: Trend analysis works well until an oil change occurs. An oil changes results in a loss of trend history and when associated by other maintenance activities, can introduce new trends such as break-in wear. *If the sample is the first sample after an oil change* or if there are insufficient samples since the last oil change to perform the linear regression calculation, calculate the trend using a *rise-over-run* equation and the time-on-oil.

$$Trend = \left(\frac{CS - Avg}{OI}\right) * SI$$

CS: current sample data,
Avg: Average of first samples after oil change,
OI: actual time on oil, and;
SI: established interval for equipment type.

Since there is no prior sample, the average data of all samples taken immediately after oil changes for this equipment type can be used to provide the missing last sample data. Alternatively, if the history prior to the oil change shows an abnormal trend, the trend data occurring before the oil change may be used to estimate or predict the trend which

should follow the oil change. Remember, destroyed trend data can never be recovered and oil change maintenance should always be scheduled with oil analysis in mind. The plot in Figure 7 shows the difference between the slope of the actual measured trend versus the slope of the predicted trend. When an abnormal condition is masked by an oil change, the data often looks like only a small amount of oil was added (Apparent Trend). In reality, the slope of the predicted trend indicates the abnormal condition has worsened, while the apparent trend slope suggests the situation is improving. While current sample data has not exceeded the established alarm, action may be required as the lost data cannot be made up without accurate knowledge of oil system additions. In most cases, trend data is more significant then level alarms.
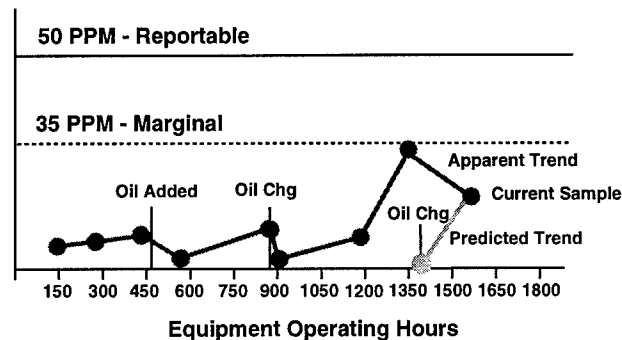

Figure 7: Calculating Trends after an Oil Change

Conclusion: Due to variability of oil test data, some form of adaptive trending is necessary for reliable condition monitoring. In addition, mathematical trending algorithms can become quite complex unless some standard method is devised. Analysis of railway, maritime and other machinery oil analysis programs has determined the above trend equations and procedures are effective at compensating for oil data variations caused by most operational and maintenance events. In these industries, autonomous oil analysis expert systems have been developed and implemented, confirming the stability and reliability of the procedures used.

Most large equipment operators, have databases with sufficient historical data to determine the critical failure modes and their indicators; reliable alarm limits; operational and maintenance events that affect data integrity or reliability; and the appropriate maintenance responses for each failure mode. With a well planned data management system, the interpretation of condition data by mathematical procedures is feasible. Moreover, the approach is consistent, thorough and reliable.